

Personalization, interaction, and navigation in rich multimedia documents for print-disabled users

H. L. Petrie
G. Weber
W. Fisher

Multimedia documents, such as textbooks, reference materials, and leisure materials, inherently use techniques that also can help make them accessible for people with disabilities who find it difficult or impossible to use printed materials. This includes individuals who are blind, partially sighted, deaf, hard of hearing, or dyslexic. The varying requirements of print-disabled users have led us to the notion of enriched media documents that contain redundant alternative representations of the same information. Unlike existing one-document-for-all approaches, we propose a personalization process that customizes these rich media documents to the needs of an individual reader. This paper describes, from an iterative user-centered design perspective, the development of a multimedia reading system for a variety of print-disabled user groups. We address issues of establishing user personalization profiles, as well as adapting and customizing content, interaction, and navigation. Customization of interaction and navigation leads to differences in the user interface, as well as different structural views of indexes. Customization of content includes insertion of a summary, synchronization of sign language video with highlighting of text, self-voicing capability, alternative support for screen readers, or reorganization of layout to accommodate large fonts. Finally, we consider whether this approach of addressing the specific needs of heterogeneous user groups provides a basis for a universal design approach for multimedia user interfaces.

INTRODUCTION

Access to text-based electronic information, whether in software applications, on the World Wide Web (WWW), or in eBooks, is a problem that has largely been solved for visually impaired individuals by using screen reading programs combined with synthetic speech or refreshable braille displays. Even if this text information is embedded in a

graphical user interface, it can be more-or-less readily extracted. However, access to truly graphic

©Copyright 2005 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the Journal reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to republish any other portion of the paper must be obtained from the Editor. 0018-8670/05/\$5.00 © 2005 IBM

and multimedia information, including images, animations, and video material, is still highly problematic for readers who are visually impaired and for readers who are deaf or hard of hearing. The solution to this accessibility problem may well lie in adoption of multimedia and format techniques already being used in electronic environments such as the WWW to provide synchronization and coordination of multimedia materials.

In the MultiReader Project funded by the EU (European Union) that is described here, we have specifically explored the use of multimedia techniques to make such material accessible to readers with a range of print-related disabilities, including visual and hearing impairments as well as dyslexia. In this paper we discuss the use of content management approaches to provide personalized multimedia documents that suit the particular needs of individual print-disabled readers, in effect creating adaptable hypermedia documents. Our central hypothesis is that the needs of all readers cannot be addressed with one single multimedia document that is transformed in different ways for each user group, the so-called "one document for all" approach. Our approach instead involves providing actual alternative media to produce a variety of different views of the document in order to address particular user subsets (*stereotypes*¹) and short-term individual preferences (the process of *personalization*).

Techniques for adaptation and personalization have not yet been applied to this area, although many ways to personalize navigation and interaction with documents for mainstream readers (without special reading needs) have been explored.² Earlier work in the AVANTI (Added Value Access to New Technologies and Services on the Internet) project addressed the needs of physically disabled and blind people by adaptation of information at the lexical, syntactic, and semantic levels of interaction.³ Access to kiosk and desktop applications was successfully provided by verbalizing the textual content through speech synthesis and replacing keyboard-based interaction techniques with single-switch operations. However, temporal relationships involving time-dependent lexical entries, such as audio or movies, were not foreseen in this application.

Addressing the needs of a stereotype requires both information about system properties suitable for a

cluster of users and information about the behaviors and actions of those users. For example, Electronic Program Guides (EPG) encode information about multiple temporal arrangements of different television genres, channels, and so forth for digital television. The television viewer interacts with an EPG and navigates through these time-dependent media. The interactive behavior of EPGs and the way that they represent the broadcast media can be modified, and alternatives based on several approaches to user modeling have been reported.⁴ However, it is important to note that the broadcast media themselves are not modified in this process. In contrast, access to multimedia documents by print-disabled individuals requires adaptation of the media themselves or provision of alternative media, with the specific nature of the changes depending on a group's particular reading needs. Alternative media may be different in their visual appearance, spatial layout, or temporal arrangement.

Considerable work has been undertaken in the context of the World Wide Web Consortium's Web Accessibility Initiative (WAI)⁵ regarding the questions of how to integrate descriptions of images into Web documents, how to provide navigation in textual documents, forms, and tables, and how to adapt the visual appearance of pages to the needs of print-disabled readers. This work has led to the concept of *media enrichment*, the providing of content in different media in addition to the original. This can involve, for example, text descriptions of images, subtitling of videos, or sign language translation of texts. Enrichment improves digital content and increases information accessibility and, at the same time, leaves the original material unchanged. These additional media used for enriching other media are in some sense redundant because they specifically do not replace the original material. However, their importance lies in the support they provide to the user in understanding the original material.

The WAI advice on accessibility for multimedia material, such as animation and video, is very general. For example, WAI recommends using alternative discrete media, but does not specify how such media might be used.⁶ In fact, for visual components of multimedia, visually impaired individuals need audio description to describe purely visual elements, whereas individuals who are deaf or hard of hearing need subtitling or sign language

translation. Currently, enrichment is usually embedded in the original medium itself for time-dependent media like video.⁷ This strict synchronization does not allow the user to *independently* control the medium used for enrichment or its presentation. Markup languages such as SMIL (Synchronized Multimedia Integration Language) are required if the document is to provide a more flexible degree of adaptation.

Multimedia documents written in SMIL can synchronize multiple time-dependent media, and even allow an author to specify temporal links among the media, whether they are presented in parallel or sequentially. This allows readers to follow the link to the beginning of a caption for a video, to hear the audio description for a movie again, or to revisit the current chapter of an audio book. The design of effective user interfaces to such documents for print-disabled users is an interesting issue because the extra time such individuals may need to operate a keyboard, scan a braille display, locate a toolbar with the mouse, or most particularly, listen to a screen reader, may conflict with other requirements, such as listening to an ongoing audio or video presentation.

This paper is organized as follows. In the following section we discuss in more detail the concept of media enrichment as well as the iterative user-centered design methodology adopted within the MultiReader project. We then describe the succession of prototypes that were developed and evaluated in the course of this project. Finally, we present our conclusions to date and suggestions for future areas of research.

MEDIA ENRICHMENT

A serious limitation of the one-document-for-all approach is that the needs of different print-disabled groups are very different. For example, a sign language video, possibly supported by additional text information, is the preferred medium of many deaf readers. Readers who are hard of hearing need to be able to select the volume of audio from background noise, human speech, or music. Dyslexic individuals require simple language or pictorial description.⁸ Elderly readers have specific requirements to make a document readable, and may need a mixture of such multimedia presentation techniques.⁹

If all of these enrichments are included in a single Web page based on HTML (Hypertext Markup

Language), they must be read by every reader.

Although assistive devices such as screen readers try to personalize Web pages by integration with the browser, in general this is an uphill battle requiring browsers to support new or novel mixtures of existing markup languages. This leads to reduced efficiency, lack of acceptance, and ultimately an unusable reading system.

Enrichment as just described includes explicit synchronization of all media using links. For example, subtitle authors need to develop their designs according to the *temporal granularity* of a video or movie. Temporal granularity of media is based on the definition of specific timing intervals by the authors. If the temporal granularity is explicitly provided, the reader will ultimately have better control over content selection. For example, color coordination of the presentation of additional media such as subtitles can be done most effectively if the subtitle author knows the timing considerations for the media in question. Audio description also needs to be synchronized, and relies in its temporal granularity on the availability of quiet gaps in the soundtrack of a video or movie.

Figure 1 shows the architecture of the MultiReader reading system. (More detail can be found in Reference 10.) Enrichment in MultiReader uses XML (Extensible Markup Language) as its framework, allowing for rich media documents in the back end of the system. The front end receives content that is adapted to a particular reader group and contains markup for personalizing the user interface to allow more effective reading. By providing sufficient redundancy and separation of content from layout, the system ensures that for any given situation adaptation is possible without involving the complete content management chain.

The development process for the MultiReader system followed the user-centered design principles described by ISO (International Organization for Standardization) Standard 13047.¹¹ In the first design iteration, we investigated navigation through the multiple views of a multimedia presentation. Our experiences showed that a second iteration was required to develop synchronization between redundant parts of the content. Finally, content and navigation were validated with multiple heterogeneous user groups in a third iteration. We describe these three steps in detail in the following sections.

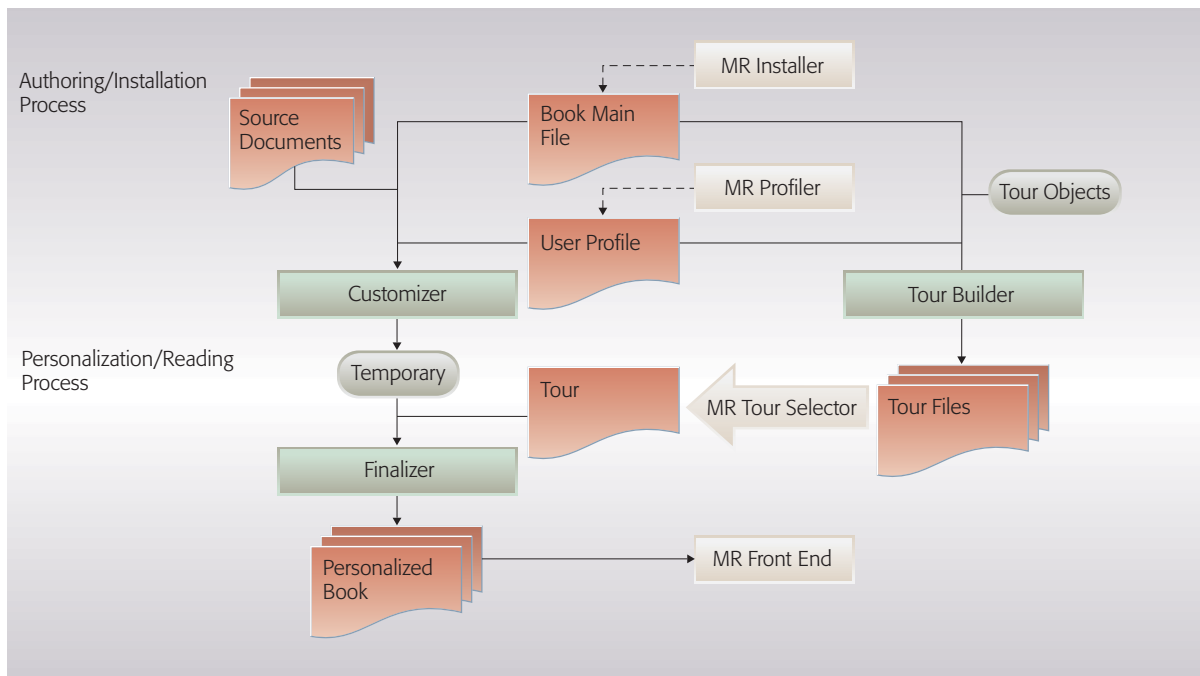


Figure 1
Architecture of the MultiReader reading system

Iteration 1: Navigation through multiple views of a multimedia presentation

To initiate the investigation of the development of enriched multimedia documents, the project developed a tourist guide for the German town of Wernigerode. This guide included text, images (including maps) coded in scalable vector graphics (SVG), and videos. Enrichments included text descriptions of the images, text annotations on the maps, and subtitling of the videos. Interface adaptations included magnification of text and maps, different text and background colors, and the highlighting of text one sentence at a time.

An evaluation with 19 partially sighted, deaf, or dyslexic users investigated the following issues, in addition to general accessibility and usability problems:

- Personalization of intradocument navigation structures (e.g., table of contents, indexes)
- Personalization of intrapage navigation structures (e.g., jump-to-top-of-page capability, location of navigation links)

Users appreciated the idea of personalization of the navigation structures, and the prototype imple-

mented this possibility with a certain degree of success. However, users suggested a number of improvements to the implementation. In particular, readers expressed the need to adjust numerous aspects of both the content and the interface to a degree that went beyond the level originally foreseen for the different stereotypes.

The development of this prototype did show that a framework of user stereotypes can provide a specification of the media enrichment alternatives that are required in the authoring of multimedia documents.

Iteration 2: Synchronization of media

For the second iteration of development and evaluation, the Wernigerode prototype was revised and extended, and an additional multimedia document, namely a tourist guide to parts of London, was also developed. Sign language videos of text were included as a further enrichment of content for profoundly deaf readers. Further adaptations of the interface included highlighting that was synchronized with the sign language, a signed navigation toolbar, and animated icons for dyslexic readers. A personalization system with three levels was developed, including the following capabilities:

1. Selection of a stereotype that addresses temporal user attributes
2. Selection of user interface elements
3. Selection of media available for enrichment and the temporal arrangement of these media

An evaluation with 12 deaf, dyslexic, and mainstream users investigated issues related to the personalization of synchronized media, again also considering general accessibility and usability problems.

The sign language videos, synchronization with highlighting, and signed navigation toolbar icons were very well received by deaf readers. (A more complete set of signed navigation icons underwent a separate quantitative evaluation and was equally well received by deaf users.¹²) The animated icons for dyslexic readers were less well received and would need further iteration to be useful. The personalization system was also very well received. Although some readers might have needed assistance in learning how to use the more detailed levels of this system to obtain optimal results, giving this degree of personalization to print-disabled readers proved very useful conceptually.

Iteration 3: Integration

For the third iteration of development and evaluation, a revised and extended version of the London tourist guide was developed. In addition both a multimedia version of Hamlet and a multimedia document on the topic of visually impaired artists were created. The latter two documents were adaptations of print documents. This allowed us to investigate the transfer of existing print materials to multimedia along with the associated implications for the authoring process.

The content enrichment provided in this iteration included audio output for the text, improved and richer descriptions of images for blind users, and signed descriptions of text for deaf users. Adaptations of the interface were now integrated into an easy-to-use profiling system that included controls for speed of highlighting and speech and for magnification of video. The navigation structures were also improved with a thematic index. For example, themes in Hamlet included murder, madness, and death.

The following additional resources were required for the production of an enriched multimedia document:

1. *Subtitling of audio and video material for deaf and hard of hearing readers.* This enrichment may also prove very useful to readers whose native language is not the language of the original material.
2. *Subtitle generation and authoring systems to support subtitling.* In fact, systems that provide at least semi-automatic subtitle generation are now available.¹³
3. *Sign language interpretation of audio and video material for profoundly deaf readers.* Sign language interpretation is preferred over subtitling by many profoundly deaf readers. Unfortunately this is a considerable additional expense if a human signer is used, as is currently preferred by most sign language users. However, virtual human avatar systems now under development present sign language in a sufficiently natural way¹⁴ that they should soon prove acceptable. This advance will greatly reduce the cost of providing sign language.
4. *Audio description of video material for visually impaired readers.* There is currently no algorithmic method available for the generation of an audio description of video.

All of these production steps could be integrated into a general multimedia authoring tool intended to support skilled editors. The overhead of producing alternative media is not great when compared to the total cost of separately producing an audio book, a large-print edition, and a sign language title on CD-ROM (Compact Disk—Read Only Memory).

An evaluation with 70 print-disabled users from all of the target user groups was carried out to investigate the full range of features developed within the MultiReader system. Both the use of sign language videos synchronized with text highlighting for deaf readers (see *Figure 2*) and text highlighting synchronized with simultaneous speech output for dyslexic readers proved particularly successful. This evaluation also showed the importance of enabling the reader to control each time-dependent medium and all time-dependent enrichment. For example, dyslexic readers need to be able to control the speed of highlighting of the text, and blind readers need to be able to start and stop videos, primarily so that

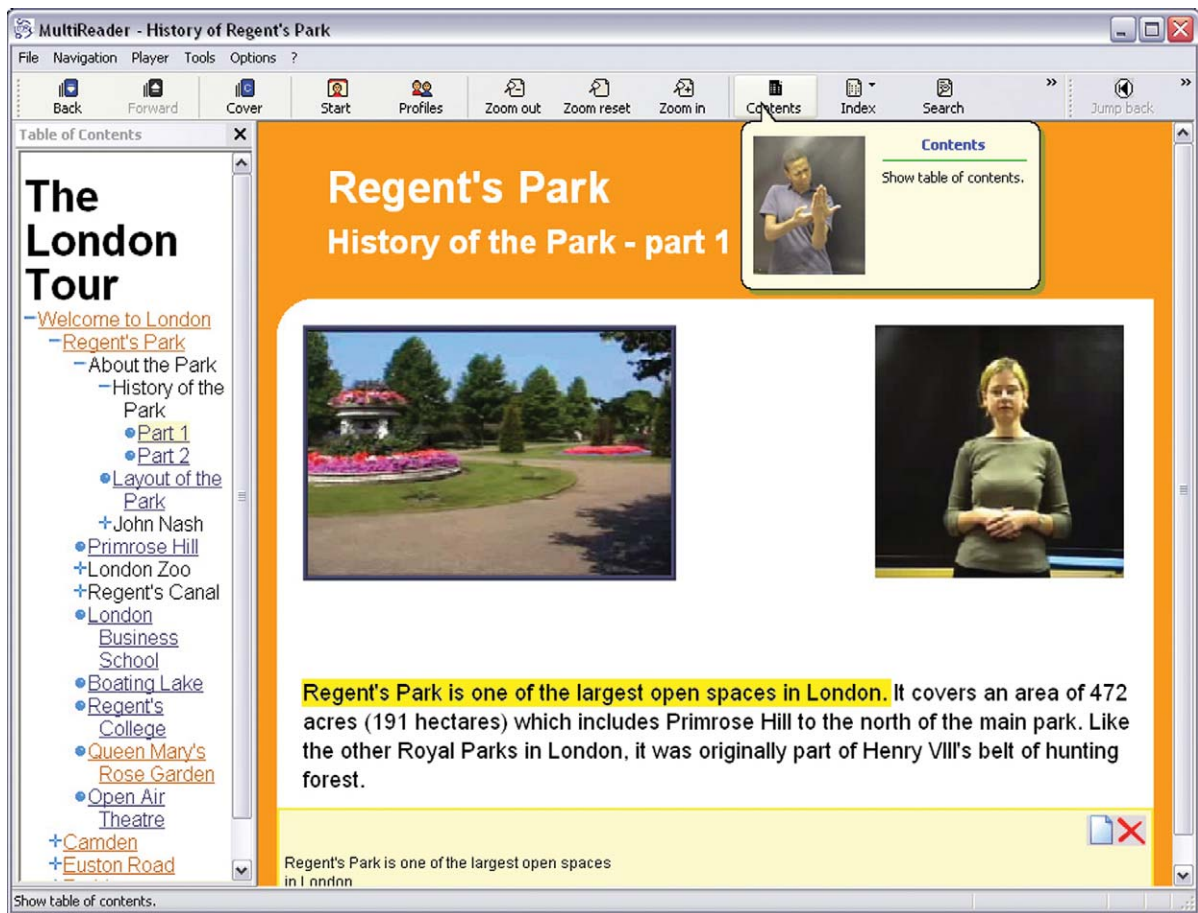


Figure 2
MultiReader as read by a deaf reader

they do not interfere with the speech from a screen reader.

CONCLUSION

Multimedia presentation can limit access to Web documents and eBooks by print-disabled readers, but multimedia documents also provide capabilities that can improve their accessibility. Existing techniques make text accessible to blind individuals, but multiple additional media such as audio description or sign language videos are needed to address the needs of deaf or dyslexic readers. Enrichment of multimedia documents not only requires adding more media but also markup describing temporal granularity. Together these give readers better control over a presentation and support them in their navigation of time-dependent media. If readers from heterogeneous user groups are involved in the iterative development of enriched multimedia Web

documents and eBooks, valuable feedback to the authors can be provided and readability ensured, despite seemingly contradicting requirements for the different reader groups. More work is required to investigate how publishers can establish quality measures for authors while supporting each of the reader stereotypes and personalization requirements.

ACKNOWLEDGMENTS

The MultiReader Project was supported under the IST (Information Society Technologies) Program by the Commission of the European Union (Project IST-2000-27513). The MultiReader Consortium consists of City University London in the United Kingdom, Katholieke Universiteit Leuven in Belgium, Royal National Institute for the Blind in the United Kingdom, Federation of Dutch Libraries for the Blind

in the Netherlands, Harz University of Applied Studies in Germany, University Kiel—Multimedia Campus in Germany, and Pragma in the Netherlands.

Transactions," *International Journal of Human-Computer Interaction* **16**, No. 2, 141–161 (2003).

Accepted for publication January 21, 2005.

Published online August 2, 2005.

CITED REFERENCES

1. E. Rich, "Stereotypes and User Modeling," in *User Models in Dialog Systems*, A. Kobsa and W. Wahlster, Editors, Springer, Berlin (1989), pp. 31–51.
2. P. Brusilovski, "Adaptive Hypermedia," *User Modeling and User-Adapted Interaction* **11**, No. 1–2, 87–110 (2001).
3. *User Interfaces for All—Concepts, Methods, and Tools*, C. Stephanidis, Editor, Lawrence Erlbaum Associates, Inc., Mahwah, NJ (2001).
4. L. Ardissono, A. Kobsa, and M. T. Maybury, *Personalized Digital Television: Targeting Programs to Individual Viewers*, Springer, Berlin (2004).
5. W. Chisholm, G. Vanderheiden, and I. Jacobs, "Web Content Accessibility Guidelines 1.0," *Interactions* **8**, No. 4, 35–54 (April 2001).
6. M.-R. Koivunen and I. Jacobs, *Accessibility Features of SMIL*, World Wide Web Consortium (September 21, 1999), <http://www.w3.org/TR/SMIL-access/>.
7. *Making Educational Software and Web Sites Accessible: Design Guidelines Including Math and Science Solutions*, CPB/WGBH National Center for Accessible Media (NCAM) (2003), <http://ncam.wgbh.org/cdrom/guideline/>.
8. H. Petrie, W. Fisher, G. Weber, I. Langer, K. Gladstone, C. Rundle, and L. Pyfers, "Universal Interfaces to Multimedia Documents," *Proceedings of 4th IEEE International Conference on Multimodal Interfaces (ICMI 2002)*, Pittsburgh, PA, October 14–16, 2002, IEEE Computer Society Press, Los Alamitos, CA (2002), pp. 319–324.
9. S. J. Czaja, "Computer Technology and the Older Adult," in *Handbook of Human-Computer Interaction (2nd Edition)*, M. Helander, T. K. Landauer, and P. Prablu, Editors, Elsevier BV, Amsterdam, The Netherlands (1997), pp. 797–812.
10. I. Langer, "Adaptation of Multimedia eBooks," in *Universal Access in HCI: Inclusive Design in the Information Society, Volume 4*, C. Stephanidis, Editor, Lawrence Erlbaum Associates, Inc., Mahwah, NJ (2003), pp. 1442–1446.
11. *ISO Standard 13407: Human-Centered Design Processes for Interactive Systems*, International Organization for Standardization, Geneva, Switzerland (1999).
12. H. Petrie, W. Fisher, K. Weimann, and G. Weber, "Augmenting Icons for Deaf Computer Users," *Proceedings of the 2004 Conference on Human Factors in Computing Systems (CHI 2004)*, Vienna, Austria, April 24–29, 2004, E. Dykstra-Erickson and M. Tscheligi, Editors, ACM, New York (2004), pp. 1131–1134.
13. A. Bateman, J. Hewitt, A. Ariyaeinia, P. Sivakumaran, and A. Lambourne, "The Quest for the Last 5%: Interfaces for Correcting Real-Time Speech-Generated Subtitles," *Proceedings of the 2000 Conference on Human Factors in Computing Systems (CHI 2000)*, April 1–6, The Hague, The Netherlands, ACM, New York (2000), pp. 129–130.
14. S. Cox, M. Lincoln, J. Tryggvason, M. Nakisa, M. Wells, M. Tutt, and S. Abbott, "The Development and Evaluation of a Speech-to-Sign Translation System to Assist

Helen L. Petrie

Centre for Human-Computer Interaction Design, City University, Northampton Square, London, EC1V0HB, United Kingdom (h.l.petrie@city.ac.uk). Professor Petrie holds the Chair in Human-Computer Interaction at City University London. She received a B.A. degree in psychology from the University of Melbourne in 1979, a Ph.D. degree in psychology from University College, London in 1988, and an M.Sc. degree in information technology from Queen Mary Westfield College London in 1992. Professor Petrie is an Associate Fellow of the British Psychological Society, a member of the British Computer Society, and a member of the Association for Computing Machinery (ACM). She has served as secretary of the Association for the Advancement of Assistive Technology in Europe (AAATE), and as vice-chair of the ACM Special Interest Group on Computing and Physical Handicaps (SIGCAPH).

Gerhard Weber

Multimedia Campus Kiel, Westring 431–453, 24118 Kiel, Germany (g.weber@mmc-kiel.com). Professor Weber holds the Chair in Human-Centered Interfaces at the Christian Albrechts University of Kiel. He received a diploma degree in computer science from the University of Stuttgart in 1984 and a doctoral degree from the University of Stuttgart in 1989. He completed his Habilitation at the University of Magdeburg in 2000. Professor Weber is a member of the Association of Computing Machinery and Gesellschaft fuer Informatik, and is vice-chair of IFIP (International Federation for Information Processing) Working Group 13.3 on HCI (human-computer interaction) and Disabilities.

Wendy Fisher

South East London Strategic Health Authority, 1 Lower Marsh, Waterloo, London, SE1 7NT, United Kingdom (wendy.fisher@selondon.nhs.uk). Ms. Fisher worked on the MultiReader Project as a researcher between 2001 and 2003. She now oversees the National Health Service Research and Development in southeast London. She received a B.Sc. degree with Honors in psychology from the University of Hertfordshire in 1999, and an M.Sc. degree in research methodology, also from Hertfordshire, in 2000. ■